Temperature dependence and counter effect of the correlations of folding rate with chain length and with native topology

Hironori K. Nakamura* and Mitsunori Takano†
Department of Physics, School of Science and Engineering, Waseda University, 3-4-1 Okubo, Shinjuku-Ku, Tokyo 169-8555, Japan
(Received 1 February 2005; published 22 June 2005)

There is a controversy about the major determinants of the folding rate of small single-domain proteins. To shed light on this issue, we examined a possibility that the major determinants may change depending on temperature by conducting molecular dynamics simulations for 17 small single-domain proteins using an off-lattice Go-like model over a wide range of temperature. It was shown that the rank order of the folding rates is temperature dependent, which indicates that the major determinants are dependent on temperature. It was also found that as temperature is decreased, the correlation of the folding rate with the chain length becomes weakened, whereas that with the native topology becomes enhanced. Our simulation results, therefore, may provide a clue to reconcile the apparent controversy between the study by Plaxco *et al.* based on experimental data and the previous theoretical and subsequent simulation studies: the former showed that the folding rate of two-state folders does not correlate with the chain length but correlates well with the native topology, whereas the latter showed that the folding rate does correlate with the chain length. We propose a possible scenario reconciling the controversy, explaining the reason why the correlation of the folding rate with the chain length became weakened and that with the native topology became enhanced with decreasing temperature.

DOI: 10.1103/PhysRevE.71.061913 PACS number(s): 87.15.Cc, 87.15.Aa, 82.39.−k

I. INTRODUCTION

Which factor determines the folding rate of a protein is one of the important questions in the protein folding problem. Plaxco *et al.* first showed that the logarithmic folding rate $\log k$ of two-state folder proteins correlates well with relative contact order (RCO) that captures the property of the backbone topology of the native conformation (i.e., "native topology"), but does not correlate with chain length (L), based on the analysis of a number of experimentally measured folding rates [1–3]. In marked contrast, theoretical studies have indicated that the folding rate does correlate with L, even though the actual L dependence appears different among theories, ranging from a power-law dependence $k \sim L^{-\nu}$ [4–7] to such an exponential dependence as $k \sim \exp(-cL^{1/2})$ [8] or $k \sim \exp(-cL^{2/3})$ [9]. By molecular dynamics (MD) simulations using an off-lattice Go-like model, Koga and Takada [10] suggested that both L and RCO are important to determine the folding rate and demonstrated that the folding rate shows the best correlation with a combination of the two, particularly in the form of $\exp(-c \times (\mathrm{RCO}) \times L^{0.6})$. In addition to RCO, other quantities that also capture the property of the native topology, such as long range-order (LRO) [11] and total contact distance (TCD) [12], have recently been proposed and $\log k$ is shown to correlate with these quantities better than with RCO, which reinforces the point that the native topology is actually one of the major determinants of the folding rate. However, another MD simulation study with an off-lattice Go-like model by Cieplak and Hoang [5] showed that there is no significant correlation between $\log k$ and RCO, arguing that correlation of the folding rate with the native topology is obscured when the native topology is measured in such an *averaged* manner as RCO.

Thus, there remains a controversy about the two candidate determinants of the folding rate: i.e., the *chain length* and the *native topology*. To shed light on this issue, we examined a possibility that the major determinants of the folding rate may change depending on temperature. Since the folding rate changes depending on temperature [13–15], as has been also demonstrated by a number of theoretical studies [4–7,16–28], the rank order of the folding rates among various proteins should be changed unless the folding rates for the proteins show the same temperature dependence. The change in the rank order of the folding rates leads to the change in the correlation coefficient of the folding rate with the candidate determinants and, hence, leads to a conclusion that the major determinants of the folding rates are changed. In preceding MD simulation studies, the correlation of the folding rate with the candidate determinants has been examined at a fixed (though scaled) temperature—e.g., at about the folding temperature [10] and at the temperature where the folding rate becomes the fastest [5]. Although there have been a number of theoretical studies focusing on the temperature dependence of folding kinetics [4–7,16–28], the temperature dependence of the correlation between the folding rate and the candidate folding-rate determinants has not yet been investigated. To this end, we conducted MD simulation for 17 small single-domain proteins using an off-lattice Go-like model, covering a wide range of temperature.

*Present address: Division of Prion Research, Center for Emerging Infectious Diseases, Gifu University, 1-1 Yanagido, Gifu 501-1194, Japan.
†Corresponding author. FAX: +81-3-5286-3512. Electronic address: mtkn@waseda.jp

II. MODEL AND METHOD

We used a course-grained off-lattice chain model consisting of only C_α atoms with the Go-like potential developed by Clementi *et al.* [29]. The so-called Go potential can be traced back to the seminal work by Go *et al.* [30]. The off-lattice Go-like models have been successfully used in the previous MD simulations of protein folding [10,21,29,31,32], and the details of the model were described elsewhere [10,21,29]. We here describe only several important points: The native contacts for a protein were determined according to the PDB structure. If any of the nonhydrogen atoms in the *i*th residue is located within 6.5 Å from any of the nonhydrogen atoms in the *j*th residue, the *i*-*j* residue pair is judged to form a native contact [10]. In this study, total 17 PDB structures were used: albumin binding domain (1PRB, 53), protein G (2GB1, 56), src-SH3 (1SRL, 56), α-spectrin SH3 domain (1SHG, 57), TN fibronection type 3 (1TEN, 89), acylphosphatase (2ACY, 98), ubiquitin (1UBQ, 76), apo-myoglobin (1PMB, 153), IM9 (1IMQ, 86), cold shock protein B (1NMG, 67), protein L (2PTL, 62), lambda repressor (1LMB, 80), cytochrome b562 (256B, 106), Sso7d (1BNZ, 64), Ada2h (1O6X, 81), CI2 (1COA, 64), and U1a (1URN, 96), where the PDB code and the chain length are shown in the parentheses. Most of the proteins are known to be two-state folders and identical with the ones used in the previous study [10].

In simulation studies, folding rate is usually calculated as the reciprocal of the mean first passage time (MFPT), which is the average of the first passage time (FPT), and the FPT is the time required for a folding trajectory to reach the native conformation for the first time. The folding rate is given by

$$k = \frac{1}{\frac{1}{n}\sum_{i=1}^{n} t_i}, \tag{1}$$

where t_i is FPT of ith folding simulation and n is the total number of simulations ($n=100$). Each folding simulation was started from an unfolded conformation, which was obtained by a presimulation at a high temperature. The folding trajectory is judged to be in the native state when the fractional number of the native contacts, Q, is larger than Q_N. Q_N was set to the value at which the free energy profile shows the minimum on the native side at the folding temperature T_f. The temperature of the MD simulation was controlled by the Berendsen thermostat [33], and T_f was determined as the temperature at which the heat capacity shows the maximum. Heat capacity as a function of temperature was calculated by the weighted histogram analysis method (WHAM) [34]. The maximum number of the time steps (t_{max}) was set to be 10^7, and the FPT was assumed to be t_{max} if a trajectory could not reach the native state within t_{max}. The folding simulations were conducted over a wide range of temperature, ranging from T_f down to $0.4T_f$ with an interval of $0.1T_f$.

We examined correlations of the folding rate with the following quantities: the chain length (L), the RCO [1], the absolute contact order (ACO[=(RCO)$\times L$]) [1], (RCO)$\times L^{0.6}$ [10], the LRO [11], and the TCD [12]. RCO, LRO, and TCD are defined as

$$(\mathrm{RCO}) = \frac{1}{N \times L}\sum_{k=1}^{N} \Delta L_{ij}, \tag{2}$$

$$(\mathrm{LRO}) = \frac{N}{L}, \tag{3}$$

$$(\mathrm{TCD}) = \frac{1}{L^2}\sum_{k=1}^{N} \Delta L_{ij}, \tag{4}$$

where N is the total number of contacts and ΔL_{ij} is the sequence separation between the residues i and j. For LRO, only the contacts with $\Delta L_{ij} > 12$ were considered in evaluating N. For the TCD, the contacts with $\Delta L_{ij} > 1$ were considered.

III. RESULTS AND DISCUSSION

A. Temperature dependence of folding rates

Figure 1 shows the logarithmic folding rate as a function of temperature for each of the 17 proteins. In this study, we refer to the logarithmic folding rate as the "folding rate." The results are shown in four panels [Figs. 1(a)–1(d)] to disentangle the 17 curves so that one can easily identify each curve. As in our previous studies [21,22], all of the proteins we studied here showed the fastest folding rate at a temperature lower than the folding temperature T_f, which is actually the property of small single-domain proteins [14,15]. However, the temperature at which the folding rate becomes fastest, $T_{k\text{-max}}$, varied from protein to protein: about half of the proteins folded fastest at around $(0.6-0.7)T_f$, whereas the other proteins folded fastest at lower temperatures. Due to the fast folder group at the lower temperatures, the rank order of the folding rates was changed depending on temperature. In particular, we note that while α-helical proteins with longer chain length folded slower than β-sheet-rich proteins with shorter chain length at about T_f, the rank order of the folding rates was reversed at lower temperatures: see, for example, the curves for 1PMB (α-helix-rich, $L=153$) and 1NMG (β-sheet-rich, $L=67$) shown in Fig. 1(a) and those for 256B (α-helix protein, $L=106$) and 1BNZ (α-sheet protein, $L=64$) shown in Fig. 1(b).

B. Folding rate determinants at T_f

We next examine the correlations of the folding rate with the candidate quantities to identify the major determinants of the folding rate at T_f. As shown in Fig. 2, the folding rate was found to correlate most strongly with ACO with $r=-0.87$ and $p<0.0001$, where r represents the correlation coefficient and p represents the probability that the null hypothesis of no correlation ($r=0$) is accepted, so that the correlation coefficient is usually considered to be statistically significant if p is smaller than 0.05. The correlation of the folding rate with L is not so strong as ACO but is substantial ($r=-0.61$, $p=0.009$), which is qualitatively consistent with the previous theoretical and simulation studies [4–10] in that L is one of the major determinants of the folding rates. In contrast, the

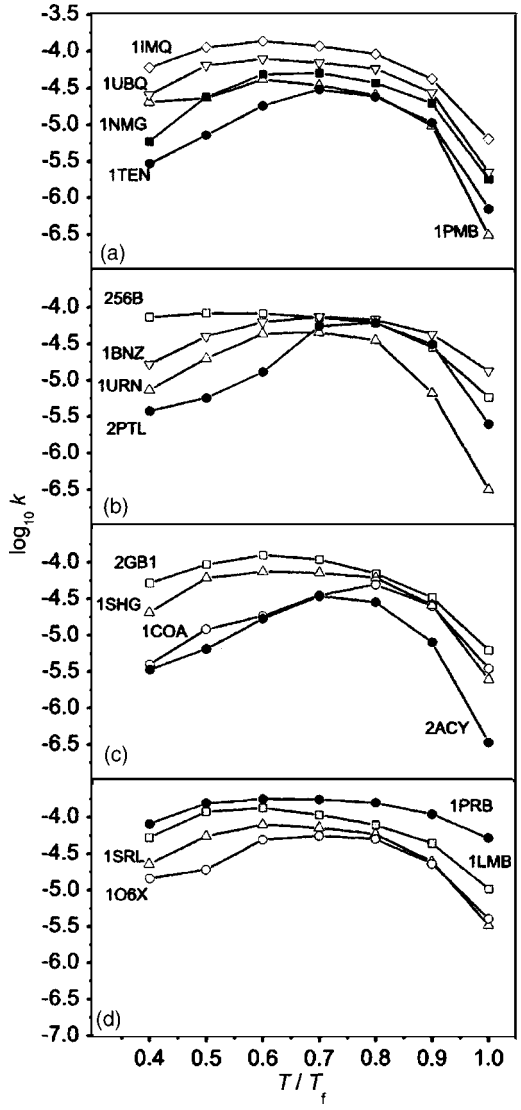

FIG. 1. Temperature dependence of the folding rates for 17 proteins. To make it easy to identify the data set for each protein, the results are divided into four panels (a)–(d). Lines are drawn for guides for eyes and labeled by PDB codes (see the text).

folding rate showed only a weak (almost negligible) correlation with RCO, which is again consistent with the previous simulation study [5].

The result we obtained at T_f, however, appears to be opposite to the one obtained by Plaxco *et al.*, which indicated that the folding rate does not correlate with L but correlates well with RCO. In this point, it is important to note that the folding rate showed a substantial correlation with another quantity, LRO, which also measures the native topology in a somewhat different way from RCO and puts more emphasis on the native contacts with larger sequence separation ($r=-0.62$, $p=0.009$). This suggests that the native topology actually plays a role in determining the folding rate. Moreover, the role of the native topology is also reflected in the fact that the folding rate correlated best with ACO: since ACO is the product of RCO, which is L invariant and purely representing the information of the chain topology, and L [i.e., $(\mathrm{ACO})=(\mathrm{RCO})\times L$], it can be interpreted that additional consideration of the pure information of the native topology represented by RCO enhanced the correlation of the folding rate with L. Conversely, it can be said that additional consideration of L enhanced the correlation of the folding rate with RCO, as seen in the result that the correlation of the folding rate with RCO became more enhanced as the involvement of L was emphasized in such a way as $(\mathrm{RCO})\times L^0$ ($r=-0.38$, $p=0.135$), $(\mathrm{RCO})\times L^{0.6}$ ($r=-0.74$, $p=0.001$), and $(\mathrm{RCO})\times L^1=(\mathrm{ACO})$ ($r=-0.87, p<0.0001$).

In the previous MD study by Koga and Takada [10], the folding rate was shown to correlate best with $(\mathrm{RCO})\times L^{0.6}$ near T_f, which is apparently different from the result we obtained; in our result, the folding rate was shown to correlate best with ACO [$=(\mathrm{RCO})\times L^1$]. This difference, however, is not significant because the difference between the correlation coefficient of $\log k$ vs ACO and that of $\log k$ vs $(\mathrm{RCO})\times L^{0.6}$ is so small that it would fall within the statistical errors. Our result, thus, is qualitatively in accordance with their indication that both the chain length and the native topology were important determinants of the folding rate.

C. Folding rate determinants at lower temperature

It is shown in Fig. 3(a) that the correlation of the folding rate with L at $0.4T_f$ ($r=-0.04$, $p=0.875$) became much

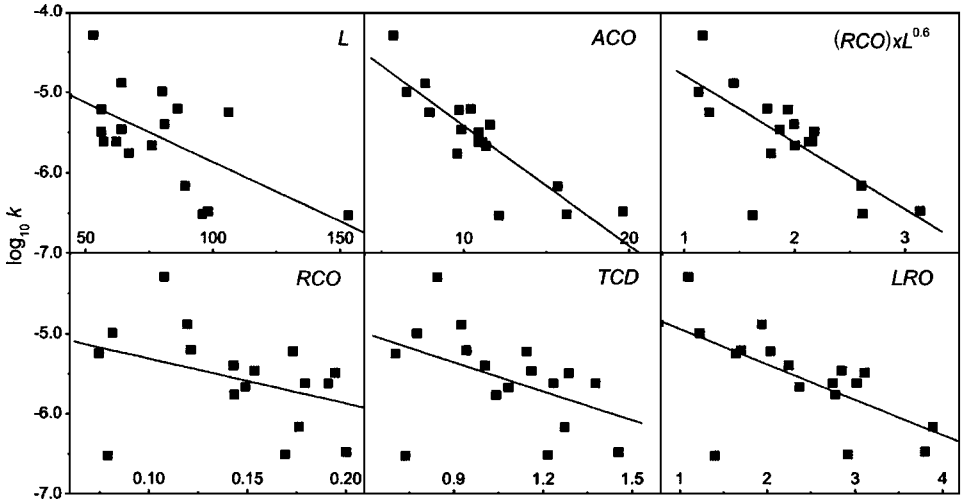

FIG. 2. Scatter plots of the folding rate at T_f against L, ACO, $(\mathrm{RCO})\times L^{0.6}$, RCO, TCD, and LRO, respectively. See text for terminologies. Lines represent the least-squares linear fits.

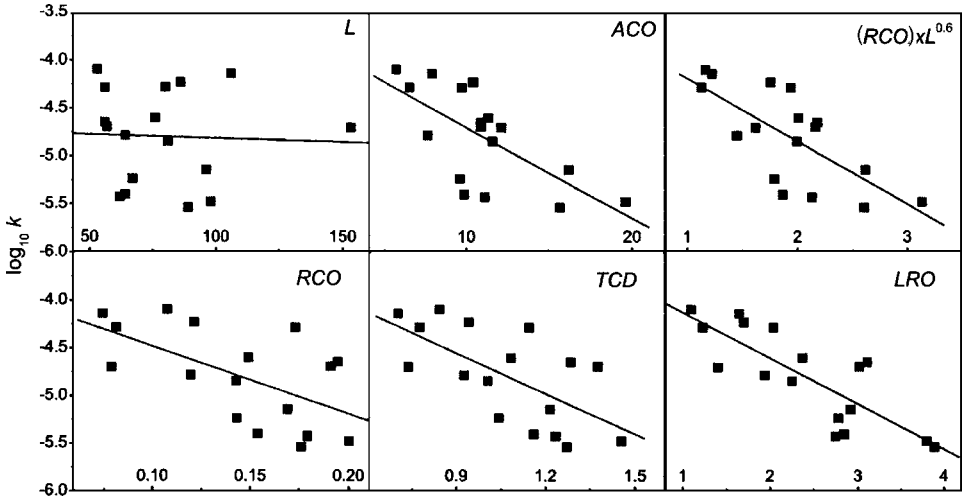

FIG. 3. The same results as shown in Fig. 2 except that the folding rates are those obtained at lower temperature ($0.4T_f$).

weaker than that found at T_f ($r=-0.61$, $p=0.009$). On the other hand, the correlation of the folding rate with RCO became enhanced ($r=-0.59$, $p=0.013$) compared to that found at T_f ($r=-0.38$, $p=0.135$) and those with the other quantities measuring the native topology also became enhanced; for LRO the correlation changed from ($r=-0.62$, $p=0.009$) to ($r=-0.81$, $p<0.009$) and for TCD from ($r=-0.45$, $p=0.069$) to ($r=-0.65$, $p=0.005$).

D. Temperature dependence of folding-rate determinants

Figure 4 shows the temperature dependence of the correlation coefficient between the folding rate and the candidate quantities, which connects the above results found at T_f and at lower temperature. It is clearly shown that as temperature was decreased, L tended to be left out of the major determinants of the folding rate whereas the quantities representing the native topology tended to be more significant determinant. This tendency is interesting because it brings our results closer to the result by Plaxco *et al.*, where the folding rate was shown not to correlate with L but correlate well with RCO [1–3]. It is also interesting that the correlation of the folding rate with the combined quantities of L and RCO—i.e., $(\mathrm{RCO})\times L^{0.6}$ and ACO—remained almost unchanged over the whole temperature range we studied. This indicates that the opposite temperature dependence of the correlation of the folding rate with the chain length and that with the native topology counterbalanced each other.

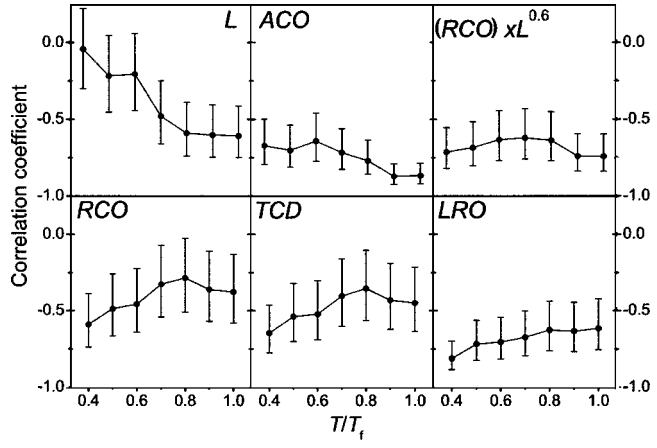

FIG. 4. Temperature dependences of the correlation coefficients between the folding rate and the quantities examined in Figs. 2 and 3. Error bars were estimated by transforming the correlation coefficient (Pearson's r value) into the Fisher's z' value, which is assumed to obey a normal distribution with the standard error of $1/\sqrt{N-3}$, where N is the number of the data.

E. Chain length as a determinant of the folding rate

We thus found that as temperature was decreased, the correlation of the folding rate with L became weakened whereas that with RCO became enhanced. To understand why the correlation of the folding rate with L became weakened with decreasing temperature, it is worth examining the pure effect of L on the folding rate. For this purpose, we analyzed the correlation of the folding rate with L separately for the proteins predominantly composed of α-helices and for the proteins predominantly composed of β-sheets. α-helix proteins form a native topology group with small RCO, whereas β-sheet proteins form another native topology group with large RCO. Figure 5(a) shows the temperature dependence of the folding rates for α-helix proteins. The curves for the proteins showed almost the same trend, running almost parallel to each other, indicating that the rank order of the folding rates among the proteins remains almost unchanged. A similar result was obtained for the β-sheet proteins [Fig. 5(b)]. Figure 5(c) shows the correlations of the folding rate with L for the α-helix proteins and for the β-sheet proteins observed at T_f, where the folding rate showed a substantial correlation with L. It is noteworthy that the folding rate showed a much better correlation with L if it was examined within each topological group. Figure 5(d) shows the corresponding results observed at the low temperature $0.4T_f$, where the folding rate did not show significant correlation with L. The folding rate again showed a strong correlation with L within each group even at this low temperature. However, in contrast with the result at T_f, the slopes of the correlation were quite different between the α-helix proteins and the β-sheet proteins, with the slope for the β-sheet proteins being much steeper than that for the α-helix proteins,

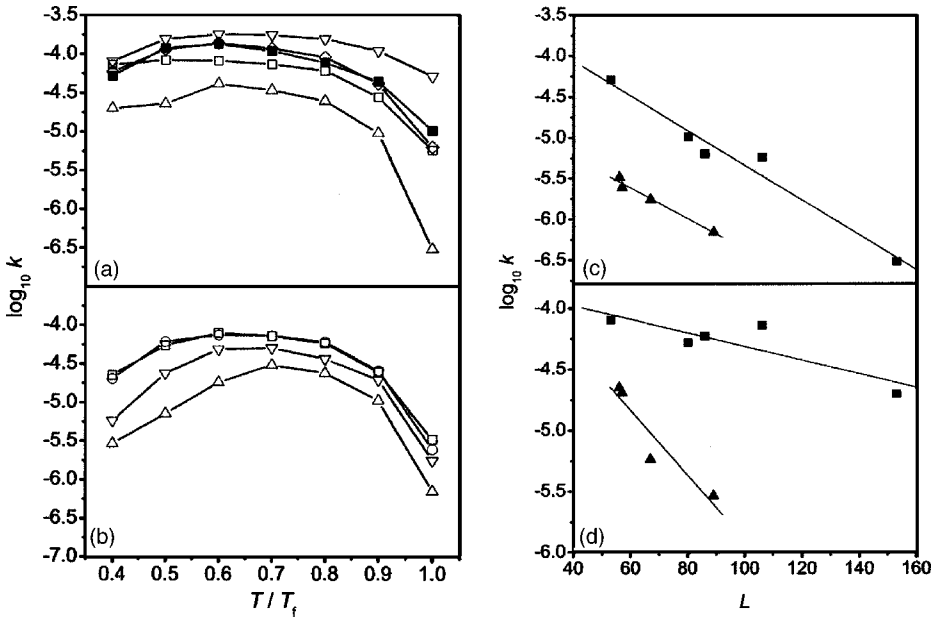

FIG. 5. Temperature dependence of the folding rate for the α-helix protein group (a) and for the β-sheet protein group (b) (data are the same as shown in Fig. 1). Plots between the folding rate and the chain length for α-helix and β-sheet protein groups observed at T_f (c) and $0.4T_f$ (d). The lines represent least-squares linear fits obtained, respectively, for the α-helix protein group and for the β-sheet protein group.

which makes the plot points considerably dispersed *as a whole*. This explains the reason why the correlation of the folding rate with L became weakened with decreasing temperature. We note that such an inherent correlation of the folding rate with L as found in our study has also been found in the previous simulation study by Cieplak and Hoang [5], in which the maximum folding rate (in logarithmic scale) was shown to correlate with $\log L$, with the slope for the β-sheet proteins being much steeper than that for the α-helix proteins. We also note that the folding rate observed in our study correlated with $\log L$, as well as with L^{λ} ($\lambda = 1/2$, $2/3$, and 1) within each topological group at each temperature (data not shown). Although discussing precise functional forms regarding the L dependence is an important issue, all we can say in this study is that L is one of the major determinants of the folding rate.

F. Native topology as a determinant of the folding rate

Figure 5 also explains the reason why the correlation of the folding rate with the native topology became enhanced with decreasing temperature. As seen in Figs. 5(c) and 5(d), the best-fit line for the β-sheet protein group is always located below that for the α-helix protein group. This means that an α-helix protein, in general, folds faster than a β-sheet protein, provided that L is fixed. However, at T_f, since the two lines are not sufficiently separated [see Fig. 5(c)], it is obvious that RCO alone cannot discriminate the folding rates; as we mentioned earlier, an α-helical protein with longer chain length folds *slower* than a β-sheet protein with shorter chain length. When the temperature was decreased down to $0.4T_f$, the two lines began to separate enough, particularly in the larger-L region. Therefore, the role of the native topology in determining the folding rate became enhanced with decreasing temperature.

G. Counter effects of chain length and native topology on the folding rate

Our simulation results thus suggest that L and RCO are both important determinants of the folding rate but their significances are changed in a temperature-dependent manner. Figure 6 shows a schematic illustration which summarizes the results: First, the folding rate correlates well with L over the whole temperature range under the condition that RCO be fixed (see the pair of light gray lines for α-helix proteins and the pair of dark gray lines for β-sheet proteins), and it also correlates well with the native topology under the condition that L be fixed (see the pair of thin lines for shorter proteins and the pair of thick lines for longer proteins). In this sense, L and RCO are thought of as inherent determinants of the folding rate. However, counter effects of these two inherent determinants on the folding rate are observed when the above conditions for L and RCO are removed and the correlations are examined against the whole set of proteins. At about T_f, the range of the folding rate for the α-helix proteins and that for the β-sheet proteins are overlapped with each other, and hence the inherent correlation of

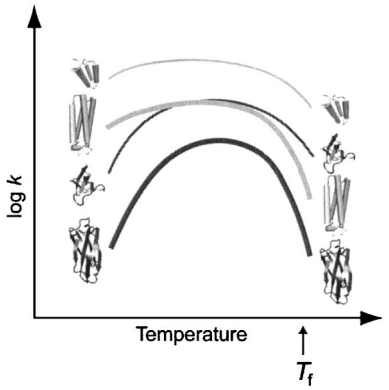

FIG. 6. Schematic illustration of the temperature dependences of the folding rate for the four typical protein groups: (i) with smaller RCO and smaller L, (ii) smaller RCO and larger L, (iii) larger RCO and smaller L, (iv) and larger RCO and larger L. The groups with smaller RCO are colored in light gray, and those with larger RCO are colored in dark gray. The groups with smaller L are indicated by thin lines, and those with larger L are indicated by thick lines.

the folding rate with RCO becomes weakened. The overlap, on the other hand, is the source of the correlation of the folding rate with L. Thus the overlap operates on L and RCO in an opposite way, and hence the effects of L and RCO on the folding rate are counter to each other. The counter effects are observed again at lower temperature in the reversed manner: the overlap is lessened compared to that observed at T_f because the folding rate for the β-sheet proteins tends to be much slower than that for the α-helix proteins. Note also that these counter effects are seen balanced as mentioned above when a combined quantity of L and RCO, such as ACO, is employed (see Fig. 4).

H. Possible scenario reconciling the controversy

The observed counter effect, as well as the temperature dependence, of the correlations of the folding rate with L and RCO, thus, may provide a clue to reconcile the controversy between the pioneering study by Plaxco *et al*. based on experimental data [1,2] and the previous theoretical and the subsequent simulation studies [4–10]: the former showed that the folding rate of two-state folders does not correlate with L but correlates well with RCO, whereas the latter showed that the folding rate does correlate with L. The lack of correlation of the folding rate with L [1–3] has been considered to be due to the lack of temperature optimization [5] or due to the narrow distribution of L [35]. In contrast, we showed that the correlation of the folding rate with L remains substantial even at temperatures other than the optimized temperature ($T_{k\text{-max}}$) and even with a narrow distribution of L. We instead showed that the correlation of the folding rate with L gets blurred in the low-temperature region below $T_{k\text{-max}}$.

It is important to point out here that experimental folding rates are typically those measured at about the room temperature (20–25 °C) [1,2], which is usually lower than $T_{k\text{-max}}$. For example, $T_{k\text{-max}}=50$ °C ($T_f=88$ °C) for cymotrypsin inhibitor 2 from barley (CI2) [14] and $T_{k\text{-max}}=42$ °C ($T_f=50$ °C) for the Engrailed Homeodomain from *Drosophila melanogaster* [15]. Therefore, we suggest that the result found on the basis of the experimental data [1–3] corresponds to our result observed at low temperature where the correlation of the folding rate with L became weakened whereas that with RCO became enhanced.

We then discuss why the effect of the native topology on the folding rate becomes enhanced in the low-temperature region. As mentioned above, the enhanced effect of the native topology on the folding rate is due to the fact that the folding rates for the β-sheet proteins tended to be much slower than that for the α-helix proteins as temperature was decreased. The energy landscape theory for protein folding [12,36,37] suggests that the folding rate is made slower as temperature is decreased because the ruggedness of the energy surface comes into effect in the lower-temperature region. In our study, because we used a Go-like model, the energy landscape should be funnel like with the energetical ruggedness suppressed. However, topological constraint, which is intrinsic to any self-avoiding chain, causes additional ruggedness of the energy surface, as we studied in the previous Monte Carlo simulation using a lattice Go model [22] (see also Ref. [28]). This kind of ruggedness caused by the topological constraint is thought to be more significant for proteins with more complicated native topology with larger RCO, which is probably the reason why the folding rates for the proteins with larger RCO tend to be much slower with decreasing temperature.

The Go-like model presents a funnel-like energy landscape with minimal frustration and, in this sense, is generally thought to follow the essential property of a real protein, particularly of a small single-domain protein. In fact, Go-like models have been successfully used in previous simulation studies of protein folding [5–7,10,21,22,27,29–32]. However, we must keep the following points in mind when comparing simulation results with experimental ones in a quantitative manner. In our simulation, $T_{k\text{-max}}$ was found to be about $0.7T_f$ or lower, which is apparently too low for a real protein. The location of $T_{k\text{-max}}$ relative to T_f, however, varies depending on the energy landscape characteristics, particularly on the ruggedness: if the ruggedness is enhanced, the folding becomes slower even at relatively high temperatures, and hence $T_{k\text{-max}}$ becomes higher. $T_{k\text{-max}}$ can even become higher than T_f, as seen in lattice model studies by Socci *et al.* [16–19] where non-native interactions are involved, whereas $T_{k\text{-max}}$ was found to be about at $0.85T_f$ in our previous study with a lattice Go model [22]. Thus it is considered that the landscape ruggedness for a real protein is underestimated in the Go-like model in which only native interactions are considered. The lack of a side chain in our Go-like model is certainly another cause of the underestimation of the ruggedness. However, incorporation of the non-native interaction and side-chain effects would not change the qualitative aspects of our results. Instead it would further enhance the role of RCO as a determinant of the folding rate in the low-temperature region. The temperature dependence of the interaction [38], which is also not taken into account in our Go-like model, is another important point when comparing with a real protein. As temperature is decreased, interactions stabilizing the native conformation are known to become weakened, leading to a reduction of the native bias in the folding funnel. This effect, too, would not change the qualitative aspects of our results, because the landscape ruggedness would become relatively more emphasized due to the reduction of the native bias. We thus expect our results are robust against these effects, although examining this expectation, of course, is an important future subject.

ACKNOWLEDGMENTS

We thank Nobuyasu Koga, Fumiko Takagi, and Shoji Takada for helpful comments and for providing their MD program. This research is supported by ACT-JST, Grant-in-Aid from MEXT, and Waseda University Grant for Special Research Projects.

[1] K. W. Plaxco, K. T. Simons, and D. Baker, J. Mol. Biol. **277**, 985 (1998).
[2] K. W. Plaxco *et al.*, Biochemistry **39**, 11177 (2000).
[3] D. N. Ivankov *et al.*, Protein Sci. **12**, 2057 (2003).
[4] A. M. Gutin, V. I. Abkevich, and E. I. Shakhnovich, Phys. Rev. Lett. **77**, 5433 (1996).
[5] M. Cieplak and T. X. Hoang, Biophys. J. **84**, 475 (2003).
[6] M. Cieplak, T. X. Hoang, and M. S. Li, Phys. Rev. Lett. **83**, 1684 (1999).
[7] M. Cieplak and T. X. Hoang, J. Biol. Phys. **26**, 273 (2000).
[8] D. Thirumalai, J. Phys. I **5**, 1457 (1995).
[9] A. V. Finkelstein and A. Y. Badretdinov, Folding Des. **2**, 115 (1997).
[10] N. Koga and S. Takada, J. Mol. Biol. **313**, 171 (2001).
[11] M. M. Gromiha and S. Selvaraj, J. Mol. Biol. **310**, 27 (2001).
[12] H. Zhou and Y. Zhou, Biophys. J. **82**, 458 (2002).
[13] A. Fersht, *Structure and Mechanism in Protein Science: A Guide to Enzyme Catalysis and Protein Folding* (Freeman, New York, 1998).
[14] M. Oliveberg, Y.-J. Tan, and A. R. Fersht, Proc. Natl. Acad. Sci. U.S.A. **92**, 8926 (1995).
[15] U. Mayor *et al.*, Proc. Natl. Acad. Sci. U.S.A. **97**, 13518 (2000).
[16] J. D. Bryngelson *et al.*, Proteins: Struct., Funct., Genet. **21**, 167 (1995).
[17] N. D. Socci and J. N. Onuchic, J. Chem. Phys. **103**, 4732 (1995).
[18] N. D. Socci, J. N. Onuchic, and P. G. Wolynes, J. Chem. Phys. **104**, 5860 (1996).
[19] N. D. Socci, H. Nymeyer, and J. N. Onuchic, Physica D **107**, 366 (1997).
[20] J. Chahine, H. Nymeyer, V. B. P. Leite, N. D. Socci, and J. N. Onuchic, Phys. Rev. Lett. **88**, 168101 (2002).
[21] H. K. Nakamura, M. Sasai, and M. Takano, Chem. Phys. **307**, 259 (2004).
[22] H. K. Nakamura, M. Sasai, and M. Takano, Proteins: Struct. Funct., Bioinfo. **55**, 99 (2004).
[23] H. K. Nakamura and M. Sasai, Proteins: Struct., Funct., Genet. **43**, 280 (2001).
[24] H. K. Nakamura, T. N. Sasaki, and M. Sasai, Chem. Phys. Lett. **347**, 247 (2001).
[25] V. I. Abkevich, A. M. Gutin, and E. I. Shakhnovich, J. Mol. Biol. **252**, 460 (1995).
[26] A. Gutin *et al.*, J. Chem. Phys. **108**, 6466 (1998).
[27] T. X. Hoang and M. Cieplak, J. Chem. Phys. **112**, 6851 (2000).
[28] M. R. Betancourt and J. N. Onuchic, J. Chem. Phys. **103**, 773 (1995).
[29] C. Clementi, H. Nymeyer, and J. N. Onuchic, J. Mol. Biol. **298**, 937 (2000).
[30] H. Abe and N. Go, Biopolymers **20**, 1013 (1981).
[31] C. Clementi, P. A. Jennings, and J. N. Onuchic, Proc. Natl. Acad. Sci. U.S.A. **97**, 5871 (2000).
[32] H. Kaya and H. S. Chan, J. Mol. Biol. **326**, 911 (2003).
[33] H. J. C. Berendsen *et al.*, J. Chem. Phys. **81**, 3684 (1984).
[34] S. Kumar *et al.*, J. Comput. Chem. **13**, 1011 (1992).
[35] K. Kamagata, M. Arai, and K. Kuwajima, J. Mol. Biol. **339**, 951 (2004).
[36] J. N. Onuchic, Z. Luthey-Schulten, and P. G. Wolynes, Annu. Rev. Phys. Chem. **48**, 545 (1997).
[37] P. G. Wolynes, J. N. Onuchic, and D. Thirumalai, Science **267**, 1619 (1995).
[38] H. S. Chan and K. A. Dill, Proteins: Struct., Funct., Genet. **30**, 2 (1998).